

ABSTRACT OF THE DISCLOSURE

Architectures based on a non-blocking fabric, such as a crosspoint switch, are attractive for use in high-speed LAN switches, ATM switches and IP routers. These fabrics, coupled with memory bandwidth limitations, dictate that queues be placed at the input of the switch. But it is well known that input-queueing can lead to low throughput, and does not allow the control of latency through the switch. This is in contrast to output-queueing, which maximizes throughput, and permits the accurate control of packet latency through scheduling.

A switch is disclosed with virtual output queueing at the input and queueing at the output. With a speedup of just four, and the use of a "most urgent packet first" method of operation, the switch can behave identically to an output-queued switch, regardless of the nature of the arriving traffic. The switch therefore performs as if it were output-queued, yet uses memory that runs more slowly.